

**APPLICATION**  
**FOR**  
**UNITED STATES LETTERS PATENT**

**INTERNATIONAL BUSINESS MACHINES CORPORATION**

## **DEVICE, SYSTEM AND METHOD FOR PREDICTIVE FAILURE ANALYSIS**

### **Field of the Invention**

- 5 This invention relates to predictive failure analysis for devices, and particularly but not exclusively to mass-produced hard disk drive devices.

### **Background of the Invention**

- 10 In the field of this invention it is known that many types of computer hardware are able to perform self-diagnosis of impending failure conditions. For example, computer hard disk drives are arranged to generate predictive failure analysis information. Furthermore error recovery in disk drives has improved in recent years such that devices can continue to function for many months or years with a high recovered error rate.

15

- However with time a point may arrive at which the unrecovered error rate becomes unacceptably high, or the number or severity of the recovered errors may become symptomatic of an impending total failure. Predictive failure analysis algorithms within the disk drive firmware are used to estimate this point and to generate alerts to users, informing them that a service action should be scheduled to replace the hardware which may be about to fail. Such alerts are critical for server based data storage systems. Although advances in redundancy and back-up technology now mean that in many cases little or no data loss will result when such a failure occurs, nevertheless the resulting 'downtime' of such a system while recovery actions are taken and new hardware is ordered and installed may be unacceptable for many business applications. Early warning of such failure enables users to plan for and minimise such disruption.

25

However, this approach has the disadvantage(s) that these predictive failure analysis algorithms are devised while the device is under development and are based upon the

data available at that time, such as early test results, experience with previous drive generations and the results of accelerated ageing tests (thermal or vibration stress). The error tolerance in such algorithms is therefore relatively wide.

- 5 It is necessary to set the threshold at which a device is called out for replacement to a fairly high level to avoid expensive hardware replacement costs, however this is difficult to do with such a wide error tolerance without risking an unacceptably high number of errors for the user.
- 10 Disk drive manufacturers typically receive predictive failure analysis information from disk drives that have been called out for replacement on a per device basis, but recovered error information is not typically received from the drives that are functioning within their tolerances.
- 15 US patent number 05123017 discloses a system in which sensors are retrofitted to different elements of a hardware system and are arranged to send information over a closed network to a central location. The information is used for diagnosing failures in order to facilitate the field replacement of faulty elements. In addition the information is used for predicting future failures.

20

A need therefore exists for an improved device, system and method for predictive failure analysis wherein the abovementioned disadvantages may be alleviated.

### **Statement of Invention**

25

In accordance with a first aspect of the invention there is provided a system for improving predictive failure attributes of distributed devices, comprising: a plurality of devices, each device including failure sensing means arranged for collecting failure

analysis data of the device and communication means coupled to the failure sensing means and arranged for transmitting the failure analysis data; a network coupled to the communication means of each of the plurality of devices; and, a server coupled to receive the failure analysis data of each of the plurality of devices via the network; wherein the  
5 server is arranged for analysing the failure analysis data received from each of the plurality of devices and for providing failure information.

In accordance with a second aspect of the invention there is provided a device comprising: failure sensing means arranged for collecting failure analysis data of the  
10 device; and, communication means coupled to the failure sensing means and arranged for transmitting the failure analysis data to a remote server via a network, wherein the server is arranged for analysing the failure analysis data received from the device and from other devices and for providing failure information.

15 In accordance with a third aspect of the invention there is provided a method for performing predictive data analysis of a number of distributed devices, the method comprising the steps of: collecting failure analysis data from a number of failure tolerant components of the number of distributed devices; transmitting the failure analysis data to a central server via a network coupled to each of the devices; processing the failure  
20 analysis data; analysing the failure analysis data received from each of the plurality of devices; and providing failure information therefrom.

Preferably the device includes an algorithm for managing the operation of the failure tolerant component and the failure information includes an updated algorithm for  
25 providing improved operation of the failure tolerant component. The updated algorithm is preferably transmitted to the device via the network.

The failure information is preferably used to improve design and manufacturing steps for future devices. Preferably it also provides an indication of operating lifespan of the devices.

- 5     Preferably the device is coupled to the network via an intermediary software agent. The intermediary software agent is preferably installed on a local server.

The local server preferably includes a database arranged for storing the failure analysis data from the device, the local server being arranged for periodically uploading the  
10     failure analysis data to the manufacturer's server.

In this way information is provided from a large population of drives in the field population, and may be used to perform detailed analysis with greater predictability and less tolerance than present arrangements. Trends or unexpected failure modes may also  
15     be detected. The information may be used to improve the operation of the hard disk drives in the field, or to make improvements to future designs.

### **Brief Description of the Drawings**

- 20     One device, system and method for predictive failure analysis incorporating the present invention will now be described, by way of example only, with reference to the accompanying drawings, in which:

FIG. 1 shows a preferred embodiment of a system for predictive failure analysis in  
25     accordance with the invention;

FIG. 2 shows a simple block diagram of a software agent forming part of the embodiment of FIG. 1.

FIG. 3 shows an alternative embodiment of a system for predictive failure analysis in accordance with the invention; and

FIG. 4 shows a flow chart of a preferred method of operation of the embodiment of FIG.

5 1.

### **Description of Preferred Embodiment(s)**

10 Referring to FIG. 1, there is shown a system 5 comprising a manufacturers server 10, utilised by a disk drive manufacturer in a manner to be further described below, coupled to a corporate Wide Area Network (WAN) 40 via the internet 20 and a firewall 30.

15 A number of disk drives 80 (manufactured by the disk drive manufacturer) are coupled to a device driver 60 via a storage adapter 70. The driver 60, adapter 70 and disk drives 80 are all operated by a customer, and used for typical disk drive applications, such as databases, servers, data storage and the like.

20 The disk drives 80 are mass-produced by the manufacturer, and may be substantially identical to each other, or may be variants of a particular family of disk drive designs. For example they may be different sizes. It is envisaged that the manufacturer may produce many thousands of disk drives of that family.

25 A software agent 50 is coupled between the device driver 60 and the corporate WAN 40. The software agent 50 may be run on a local server 55 which attaches to the disk drives 80, or by other means. The software agent 50 gathers recovered error and other predictive statistical data from the disk drives 80, and in a preferred embodiment this predictive failure data is temporarily stored in a local database 57 on the local server 55.

Periodically the predictive failure information is uploaded to the manufacturers server 10 via the corporate WAN 40 and internet 20.

5 It will be appreciated that the data can be directly uploaded to the manufacturers server 10 as it becomes available. However the database 57 helps to address the scaling issue as the manufacturers server 10 may be servicing field populations numbering many thousands or millions of individual devices.

10 The protocol used to upload data to the manufactures server 10 is not central to the invention but should be selected so that it can easily pass through the fire-wall 30 between the corporate WAN 40 and the Internet 20. In the preferred embodiment an http protocol is used because most corporate fire-walls are able to pass http requests.

15 If the device driver 60 of the disk drives 80 is connected directly to a network such as the corporate WAN 240 which supports a TCPIP protocol then the driver 60 itself could connect to the manufacturers server. However if, as described above, the driver 60 is connected to a different type of network such as a SCSI (Small Computer Systems Interface) bus then an intermediary software agent such as the software agent 50 described above is required.

20

Referring now also to FIG. 2, there is shown the internal structure of the software agent 50. A disk interrogator 110 uses SCSI commands to interrogate the attached disk drives 80. The gathered data is placed into the local database 57 by element 120. Periodically, a http client 130 connects to the manufacturers server 10 to relay the data.

25

Referring now also to FIG. 3, there is shown an alternate embodiment of the invention which incorporates a Storage Area Network (SAN) 290 as follows. Manufacturers server 210, corporate WAN 240, internet 220 and firewall 230 are identical to their counterparts in FIG. 1.

The disk drives 280 are coupled to the SAN 290, which in turn is coupled to device driver 260 via storage adapter 270. The driver 260, adapter 270, disk drives 280 and SAN 290 are all operated by a customer, and used for typical disk drive applications such as databases, servers, data storage and the like.

Software agent 250 is coupled to the disk drives 280 via the SAN 290, using a path other than that used for normal I/O operations with respect to the disk drives 280. In this way the predictive failure data is sent to the manufacturers server 210.

Referring now also to FIG. 4, there is shown an illustrative flow diagram of the capture and use of predictive failure data, according to the embodiments above. Data from disk drives typically distributed around the world are gathered locally by their respective software agents (block 300), and may be temporarily stored in local databases (block 310) or sent directly to the manufacturers server (block 320).

When the predictive failure data reaches the manufacturers server it is processed (block 330) and the results may be used for a number of purposes:

Firstly, (block 350) new microcode with improved error recovery may be made available for existing drives, targeted at the unexpected failure mode. Also new microcode may be provided which is more tolerant of certain error events than the original drive microcode and which will not call for unnecessary early drive replacement, where it has been established that the original algorithm was too aggressive, thus reducing service cost.

In both of these cases the new microcode may be available for download via the internet 20/220, or may be sent from the manufacturers server 10/210 directly to the software agent 50/150 (and to each software agent of the field population of disk drives. In this



way the predictive failure analysis algorithms of each disk drive in the field may be continually improved rather than being fixed from the date of manufacture.

Secondly, (block 360) a detected failure mode may be used to provide design changes in the microcode or manufacturing methods for new drives, so as to reduce the likelihood of the detected failure mode occurring in the future.

Finally, (block 370) planning and budgeting considerations may be made by the manufacturer for increased or decreased drive replacement if trends in the data show that the drive population is ageing faster or slower than was predicted.

It will be understood that the device, system and method for predictive failure analysis described above provides the following advantages:

Recovered error information is provided from a large population of drives in the field population, and may be used to perform detailed analysis with greater predictability and less tolerance than present arrangements. Trends or unexpected failure modes may also be detected.

It will be appreciated by a person skilled in the art that alternative embodiments to those described above are possible. For example the above invention is applicable to a wide range of mass produced devices which currently are or may be in the future connected to a network including computer tape drives, printers, automobile engine management computers, mobile phones, washing machines and the like.

Furthermore it will be understood that the means of exchanging data between the disk drives 80/280 and the manufacturers server 10/210 may differ from that described above. For example for a disk drive not coupled to the internet, removable storage media may be used for the exchange of data. Similarly, disk drives which are coupled directly to the

manufacturers server 10/210 (by a peer-to-peer arrangement or by virtue of being on the manufacturers network) do not need to use the internet 20/220.